

<https://helda.helsinki.fi>

Plasma Proteomics Analysis Reveals Dysregulation of Complement Proteins and Inflammation in Acquired Obesity-A Study on Rare BMI-Discordant Monozygotic Twin Pairs

Sahebekhtiari, Navid

2019-07

Sahebekhtiari , N , Saraswat , M , Joenväärä , S , Jokinen , R , Lovric , A , Kaye , S , Mardinoglu , A , Rissanen , A , Kaprio , J , Renkonen , R & Pietiläinen , K H 2019 , ' Plasma Proteomics Analysis Reveals Dysregulation of Complement Proteins and Inflammation in Acquired Obesity-A Study on Rare BMI-Discordant Monozygotic Twin Pairs ' , Proteomics. Clinical applications , vol. 13 , no. 4 , 1800173 . <https://doi.org/10.1002/prca.201800173>

<http://hdl.handle.net/10138/312945>

<https://doi.org/10.1002/prca.201800173>

acceptedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.

PLASMA PROTEOMICS ANALYSIS REVEALS DYSREGULATION OF COMPLEMENT PROTEINS AND INFLAMMATION IN ACQUIRED OBESITY – A STUDY ON RARE BMI-DISCORDANT MONOZYGOTIC TWIN PAIRS

Navid Sahebekhtiari^{1,*†}, Mayank Saraswat^{2,3,*}, Sakari Joenväärä^{2,3}, Riikka Jokinen¹, Alen Lovric⁴, Sanna Kaye¹, Adil Mardinoglu^{4,5,6}, Aila Rissanen¹, Jaakko Kaprio^{7,8}, Risto Renkonen^{2,3§}, Kirsi H. Pietiläinen^{1,9§†}

1. Obesity Research Unit, Research Programs Unit, Diabetes and Obesity, University of Helsinki, Helsinki, Finland
 2. Transplantation Laboratory, Faculty of Medicine, University of Helsinki, Helsinki, Finland
 3. HUSLAB, Helsinki University Hospital, Helsinki, Finland
 4. Science for Life Laboratory, KTH-Royal Institute of Technology, Stockholm, Sweden
 5. Department of Biology and Biological Engineering, Chalmers University of Technology, Gothenburg, Sweden
 6. Centre for Host–Microbiome Interactions, Dental Institute, King's College London, London, United Kingdom
 7. Department of Public Health, Finnish Twin Cohort Study, University of Helsinki, Helsinki, Finland
 8. Institute for Molecular Medicine Finland, FIMM, University of Helsinki, Helsinki, Finland
 9. Abdominal Center, Endocrinology, Helsinki University Central Hospital and University of Helsinki
- ^{*}, [§] Both authors contributed equally

† Corresponding authors: Navid Sahebekhtiari, MSc, PhD, Kirsi Pietiläinen, MD, PhD

navid.sahebekhtiari@helsinki.fi, kirsi.pietilainen@helsinki.fi

Diabetes and Obesity Research Program, P.O.Box 63, Haartmaninkatu 8, FI-00014 University of Helsinki, Finland, FAX + 358 2941 26382

Abbreviations: BMI (Body Mass Index), CRP (C reactive protein), DAVID (The Database for Annotation, Visualization and Integrated Discovery), HDL (High density lipoprotein), LDL (Low density lipoprotein), MZ (Monozygotic), MRI (Magnetic resonance imaging)

Keywords: *acquired Obesity, complement cascade, monozygotic twins, label-free proteomics, plasma proteomics*

Words: (Main text and references: 6481, including tables: 8745)

Clinical relevance

Obesity can lead to a myriad of health complications, including non-alcoholic fatty liver disease, low-grade inflammation and insulin resistance. Yet the early molecular mechanisms that underlie the development of the health complications remain incompletely understood, and studies of human obesity are often confounded by the complexity of contributing genetic and environmental factors. Here, we aimed to elucidate the early alterations in the plasma proteome associated with excess body weight independently of genetic factors, by studying a unique cohort of monozygotic twin pairs

Received: 10 18, 2018; Revised: 12 27, 2018; Accepted: 01 24, 2019

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1002/prca.201800173](#).

This article is protected by copyright. All rights reserved.

discordant for BMI. Subgroup analysis of pairs discordant and concordant for liver fat was used to assess the effect of liver fat, indicative of progression towards metabolic complications. Our plasma proteomics data demonstrated that aberrant complement activation and inflammation associate with excess body weight in clinically healthy individuals, independently of genetic factors or liver fat. The findings suggest that systemic complement dysregulation and inflammation, previously linked to obesity-related complications, are very early events in obesity. Thus, countering them could be protective against metabolic complications and help preserve metabolic health.

Abstract

Purpose. The purpose of this study was to elucidate the effect of excess body weight and liver fat on the plasma proteome without interference from genetic variation.

Experimental design. The effect of excess body weight was assessed in young, healthy monozygotic twins from pairs discordant for body mass index (intrapair difference (Δ) in BMI > 3 kg/m², n=26) with untargeted LC-MS proteomics quantification. The effect of liver fat was interrogated via subgroup analysis of the BMI-discordant twin cohort: liver fat discordant pairs (Δ liver fat > 2%, n=12) and liver fat concordant pairs (Δ liver fat < 2%, n=14), measured by magnetic resonance spectroscopy.

Results. 75 proteins were differentially expressed within the BMI-discordant pairs, with significant enrichment for complement and inflammatory response pathways in the heavier co-twins. The complement dysregulation was found in obesity in both the liver fat subgroups. The complement and inflammatory proteins were significantly associated with adiposity measures, insulin resistance and impaired lipids.

Conclusions and clinical relevance. The early pathophysiological mechanisms in obesity are incompletely understood. We showed that aberrant complement regulation in plasma is present in very early stages of clinically healthy obese persons, independently of liver fat and in the absence of genetic variation that typically confounds human studies.

Introduction

Obesity is an intractable major public health burden, affecting, along with overweight, close to 40% of the world's population ^[1]. While obesity is frequently associated with metabolic abnormalities, such as insulin resistance, type 2 diabetes and cardiovascular diseases ^[2], development of the concomitant diseases varies widely, and a notable subgroup of subjects with obesity especially during younger ages appears to be metabolically healthy ^{[3] [4]}. Even though the detailed mechanisms underlying such phenomena are not fully understood, some data suggest that ectopic fat deposition to the liver is a contributing factor to the development of metabolic abnormalities, since healthier and insulin sensitive individuals, despite obesity, have less liver fat ^{[5] [6] [7] [8]}. Thus, improved characterization of the early molecular alterations in obesity, which may

contribute to the metabolic derangements, perhaps through liver fat accumulation, is pivotal for facilitating the development of more effective, targeted approaches for their treatment and prevention.

Elucidation of underlying molecular pathways requires global quantitative methods. MS-based proteomics is ideal for identification and quantification of protein level differences in a system and offer a detailed characterization of metabolism in healthy versus diseased states. As proteins provide a broad picture of patient phenotype^{[9] [10]}, proteomics discovery approach can be used to discover the molecular malfunctions related to diseases and their progression in individual patients. Accessibility and the extensive potential to reflect metabolic state^{[9] [11] [12] [13]} makes human blood plasma and serum the predominant source material for clinical analyses. However, there appears to be a surprising paucity of data in the literature characterizing the effect of obesity on the global proteome of human plasma. This could be in part explained through the high dynamic range of plasma, which impedes the identification of more than a few hundred of the most abundant proteins by liquid chromatography mass spectrometry (LC-MS). However, depletion of highly abundant plasma proteins can partially overcome this issue^[14]. The few published studies reported variable global protein alterations in plasma or peripheral blood mononuclear cells of individuals with obesity, and mostly highlighted proteins of either the complement cascade or immune system as upregulated in obesity^{[15] [16] [17] [18]}.

Genetics, multiple environmental factors, age and sex may play a role in an individual's metabolic and plasma proteomic response to obesity. Indeed, human plasma proteomic signatures can be highly variable between individuals^{[14] [18]}. To this end, we used in the present study a rare monozygotic (MZ) twin-pair design where the co-twins are discordant for body mass index (BMI), which uniquely enables the control for genetic background, early development and environment, age and sex between heavy and lean individuals. We characterized obesity-related human plasma proteomic changes through MS-based proteomics analysis, after depleting the 12 highly abundant plasma proteins. First, we compared the proteomics profiles within all the BMI-discordant twin pairs and then, to assess the particular role of ectopic accumulation of fat to the liver on plasma proteomics, we further divided the twin pairs into those who were discordant or concordant for liver fat and analyzed the within-pair proteomics differences in these subgroups. We hypothesized that BMI-discordance with liver fat discordance results in a metabolically more disadvantaged plasma proteome than BMI-discordance without liver fat discordance. We also analyzed the overall associations between plasma proteins and other adiposity measurements, lipids and insulin

resistance, with the aim to find the proteins most closely correlating with metabolic disturbances in obesity.

Materials and methods

Participants

The present study included 26 rare monozygotic (MZ) pairs discordant for BMI (within-pair difference, $\Delta \text{BMI} \geq 3 \text{ kg/m}^2$), identified from two population-based twin cohorts, FinnTwin16 ($n = 2839$ pairs) and FinnTwin12 ($n = 2578$ pairs) ^[19]. One obese co-twin had type 2 diabetes and used metformin and insulin. Another obese co-twin had inactive ulcerative colitis and used mesalazine and azathioprine. All other participants were healthy, normotensive and did not use any medications except for oral contraceptives. Detailed descriptions of the twin material have been previously published ^{[8] [20] [21]}. Written informed consent was obtained from all participants. The study protocols were approved by the Ethical Committee of the Helsinki University Central Hospital (DNRO 270/13/03/01/2008).

Clinical measurements

Weight and height were measured for the calculation of body mass index (BMI). Body composition was determined by dual-energy X-ray absorptiometry (Lunar Prodigy, Madison, WI, software version 8.8) ^[22], the amount of subcutaneous and intra-abdominal adipose tissue by magnetic resonance imaging (MRI) ^[23] and liver fat by proton magnetic resonance spectroscopy with a 1.5 Tesla MRI imager (Avanto, Siemens, Erlangen Germany) ^[20].

Dietary intake was assessed from 3-day food records and analyzed by the Diet32 program (Aivo), based on a national Finnish database for food composition (Fineli, www.fineli.fi, National Institute for Health and Welfare, Nutrition Unit, Helsinki, Finland). Information on weekly alcohol intake during the past 4 weeks was obtained by structured questionnaires and total physical activity by the Baecke questionnaire and transformed into number of standard drinks (12 grams of alcohol per drink) per week and total physical activity by the Baecke questionnaire ^[24].

Analytical measurements

The concentrations of fasting plasma glucose, insulin during a 2-hour oral glucose tolerance test, and homeostatic model assessment (HOMA)-insulin resistance and Matsuda insulin sensitivity indexes

were calculated, as previously described ^[25]. Fasting serum high sensitivity C-reactive protein (hsCRP) was measured by Cobas CRP [Latex]HS (Roche Diagnostics). Fasting plasma total cholesterol, high-density lipoprotein (HDL) and low-density lipoprotein (LDL) cholesterol and triglyceride concentrations were determined with enzymatic methods (Abbott Architect c16000 and c8000, Abbott).

Plasma proteomics

Plasma samples from EDTA tubes, obtained after centrifugation for 15 min at $380 \times g$, were used for proteomics analyses. Multiple affinity removal columns designed to remove high abundant plasma proteins. The workflow is described in detail ^[26]. Briefly, top 12 high-abundant proteins were depleted by TOP12 protein depletion kit (Pierce, ThermoFisher) according to the manufacturer's instructions. In order to make sure equal amounts of proteins from all samples were processed for digestion and to ensure downstream equal loading onto the LC-MS, the TOP 12 proteins depleted plasma was assayed by BCA assay kit (Pierce, Thermo Scientific, Rockford, IL, USA) for the total protein concentration. The depleted proteins were reduced and alkylated by incubating samples first in 10 mM of DTT for 60 minutes at room temperature (RT) followed by incubation in 40 mM iodoacetamide for 60 minutes in the dark at RT. Excess iodoacetamide was quenched by adding 40mM DTT for 1 hour at RT. Later samples were diluted 1:10 with ultrapure water (Milli-Q, EMD Millipore corp.) and digested with trypsin (1:50 w:w trypsin to protein ratio) for overnight at +37°C. Samples were cleaned by Pierce C18 columns (Pierce, ThermoFisher) according to the manufactures protocol.

LC-MS and quantification

Samples equivalent to ~1.4 µg total protein were injected to nano Acquity UPLC (Ultra Performance Liquid Chromatography) - system (Waters Corporation, MA, USA). TRIZAIC nanoTile 85 µm × 100 mm HSS-T3u wTRAP was used in the LC system before mass spectrometer. Samples were loaded, trapped and washed for two minutes with 8.0 µL/min with 1% B. The analytical gradient was as follows:

0–1 min 1% B, at 2 min 5% B, at 65 min 30% B, at 78 min 50% B, at 80 min 85% B, at 83 min 85% B, at 84 min 1% B and at 90 min 1% B with 450 nL/min. Buffer A: 0.1% formic acid in water and Buffer B: 0.1% formic acid in acetonitrile.

DIA (data independent acquisition using HDMSE mode with Synapt G2-S HDMS (Waters Corporation, MA, USA) was used for data acquisition. The collected data range was 100–2000 m/z, scan time one-second, IMS wave velocity 650 m/s and collision energy ramp was employed in trap between 20 and 60 V. Calibration was done with Glu1-Fibrinopeptide B MS2 fragments and as a lock mass, Glu1-Fibrinopeptide B precursor ion was used during the runs. Data was collected in triplicates for every sample and further analysis was done with, Progenesis Q1 for Proteomics – software (Nonlinear Dynamics, Newcastle, UK).

Data analysis was performed essentially as previously described ^[26]. Briefly, the raw files were imported to Progenesis Q1 for proteomics software (Nonlinear Dynamics, Newcastle, UK). Lock mass correction with 785.8426 m/z, corresponding to doubly charged Glu1-Fibrinopeptide B was employed. Parameters for peak picking and alignment algorithm were default in the software. The software performed the peptide identification with Protein Lynx Global Server and label-free quantification ^[27] was used. The peptide identification was done against Uniprot human FASTA sequences (UniprotKB Release 2015_09, 20205 sequence entries) with peptides of CLPB from E. Coli (CLPB_ECOLI (P63285)) as an internal standard, ClpB protein sequence appended to FASTA file for label-free quantification. Equal amount of pre-known mixture of CLPB peptides was added to each sample and post MS acquisition was used to normalize the raw intensities. Fixed modification at cysteine (carbamidomethyl) and variable at methionine (oxidation) was used. Trypsin was specified as digesting agent and one missed cleavage was allowed. Fragment and peptide error tolerances were set to auto and FDR to less than 4%. This is the maximum allowed FDR and in these settings the search stops when FDR reaches 4%. All peptides where the error was more than 10 ppm were deleted and not considered further for analysis. Minimum one fragment per peptide, three or more fragments per protein and one or more peptides per protein were required for ion matching which also happen to be default parameters. The identified proteins are grouped using parsimony principle and peptides unique to the protein are also reported. However, a strict parsimonious approach is not taken and more details can be found at the software website (www.nonlinear.com).

Statistical analyses

The statistical analyses were performed using Stata statistical software (release 15.0; Stata Corporation), R language and environment for statistical computing (version 3.4.1, R Core Team) and Python programming language (version 3.6.5, Python Software Foundation). Comparisons of the heavier vs leaner co-twins' clinical measurements in the BMI-discordant groups were analyzed by paired t-tests and comparisons of the unrelated lean-lean and heavy-heavy co-twins and Δ values

between pairs of liver fat discordant vs. liver fat concordant subgroups by unpaired t-tests. Differential proteomics analysis between co-twin groups were conducted on log10 normalized data using empirical Bayes approach presented in LIMMA^[28], with default settings. In short, the method relies on the analysis of the expression matrix where each row represents genomic feature, and each column corresponds to a sample of the given study. The linear modeling is performed on the expression matrix in the row-wise fashion, with both coefficients and standard errors for the comparison of interest. Once linear model is fitted coefficients are used to compute statistics for all genomic features using empirical Bayes moderation of the standard errors^[29]. This approach was shown to be more robust compared to standard T-test, especially in settings with a small sample size^[30]. We encourage interested users to read more detailed statistical discussion about empirical Bayes approach used by LIMMA elsewhere^[28]. Correlation between proteins and clinical variables summed over all twin individuals (from concordant and discordant pairs) were calculated by using Pearson correlation, corrected for age and sex and family relationship within the twin pairs as described by Price *et al.*^[31]. All p values are reported after FDR correction (Benjamin-Hochberg) with significance level of 0.05.

Pathway analyses

Three different pathway analyses were performed for differentially expressed proteins (nominal p value < 0.05) between the co-twins. First, after the Gene ontology (GO) annotation for proteins detected was retrieved, the Database for Annotation, Visualization and Integrated Discovery (DAVID) was used for functional annotation analysis including GO enrichment. Second, the Ingenuity Pathway Analysis (IPA), (QIAGEN Inc., <https://www.qiagenbioinformatics.com/products/ingenuitypathway-analysis>) was applied. Third, the same dataset was applied to STRING version 10.5 (<http://string-db.org>) in order to elucidate more details about protein–protein interactions and protein networks.

Data availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE^[32] partner repository with the dataset identifier PXD010691.

Results and discussion

Clinical characteristics

The detailed clinical characteristics of the twins are presented in Table 1. In brief, and as published before^[33]^[34], the heavier co-twins of the BMI-discordant pairs (n= 26) weighed on

average 18 kg more, had significantly higher fat percentage and more subcutaneous, intra-abdominal and liver fat, were more insulin resistant, and had higher LDL cholesterol and triglyceride and lower HDL cholesterol levels than the lean co-twins. The BMI-discordant twin pairs were then further divided into liver fat discordant (n=12) and liver fat concordant (n=14) subgroups (median Δ liver fat cut-off 2%), as described earlier^{[8] [35]} (Figure S1). These subgroups were equally discordant for overall fatness, but the liver fat discordant group had significantly larger differences in intra-abdominal fat, LDL cholesterol and hsCRP within pairs than the liver fat concordant group (Table 1). Also, the heavy co-twins of the liver fat discordant subgroup had significantly higher levels of liver fat, intra-abdominal adipose tissue and insulin resistance than the heavy co-twins of the liver fat concordant group. In lifestyle questionnaires, the heavy co-twins of the liver fat discordant group reported more alcohol consumption and less physical activity than their lean counterparts, which may have contributed to the higher liver fat content. No significant lifestyle differences were observed within the liver fat concordant group.

Comparison of plasma proteome profiles between the co-twins

To elucidate the biological impact of acquired excess body weight on the plasma proteome, we first compared the co-twins in all of the 26 BMI-discordant MZ twin pairs. Across the individuals, we identified and quantified 469 plasma proteins, of which 75 proteins were significantly altered within the BMI-discordant pairs (nominal $p < 0.05$) (Table 2). Of these, 19 were downregulated and 56 were upregulated in the heavy compared with lean co-twins. Many of the identified proteins are previously reported markers for obesity, e.g. increased circulating levels of apolipoprotein B-100 (APOB)^[36], complement component 3 (C3)^[37], and serum amyloid A1 protein^{[18] [38] [39]}, validating our study model and method.

Next, we investigated the contribution of acquired liver fat on the plasma proteome by analyzing within-pair differences in the liver fat discordant and concordant groups separately. In the liver fat discordant pairs, we detected 31 significantly altered proteins; of these 12 were downregulated and 19 upregulated in the heavy as compared with lean co-twins. In the liver fat concordant pairs, 64 proteins were significantly altered, 15 and 49 were down- and upregulated, respectively, in the heavy co-twins (Tables S1- S4). It is perhaps surprising that more significantly altered proteins were detected in the liver fat concordant subgroup. As liver fat accumulation has been previously shown to be a marker of progression towards metabolically unhealthy obesity, we hypothesized the liver fat discordant twins to differ more significantly for plasma proteome profiles.

However, based on the results of our study, it does not seem that liver fat discordance produces larger differences to the plasma proteome than liver fat concordance, as we hypothesized initially.

There were 5 common proteins significantly differentially expressed between the co-twins in both the liver fat subgroups (Figure 1A, B), all of which were up-regulated in the heavy co-twins (Figure 1B). Interestingly, these were up-regulated in the entire group of BMI-discordant pairs (n=26) as well. This may suggest their putative function in obesity in general, at very early obesity stages and irrespective of the liver fat status. Of these 5 proteins, APOB is the primary protein in LDL cholesterol, a key feature of dyslipidemia and cardiovascular disease ^[36]. Coagulation factor VII has been found to be increased in obesity and metabolic syndrome previously ^{[40] [41]}. Complement component 4B (C4B) and complement factor H-related protein 5 (CFHR5) are proteins from the complement pathway of the innate immune system, and C4B was reported as increased in the serum of obese individuals ^[16]. Potential link of nuclear pore complex protein Nup107 to obesity is at present unknown. However, its' circulating levels have previously been upregulated in severe infections ^[42].

Pathway analysis of the significantly altered proteins between the co-twins

To construct an overall understanding of how the differentially expressed proteins are linked to each other, and to identify enriched functional annotation terms within the datasets, we applied protein pathway analysis. First, all the 75 statistically significantly modified proteins in plasma from all BMI-discordant co-twins were analyzed in DAVID, which identified 18 significant clusters in GO (p value < 0.05, enrichment score > 1.3). Interestingly, all of the top 5 biological processes were related to the immune system. The complement system was particularly over-represented (Figure 2A), and based on the direction of differential expression, overactivated in the heavy co-twins. In line with our DAVID analysis, the complement system and inflammation were among the top 5 enriched canonical pathways in the IPA (Figure 2B) and in the STRING pathway analyses (Figure 2C). The complement system has been linked to either obesity or weight loss in previous serum proteomics studies as well ^{[16] [18] [43]}.

STRING, DAVID, and IPA pathway analysis within liver fat concordant subgroup also identified complement activation and humoral immune response (Figure S2). Despite the fewer enriched pathways in liver fat discordant subgroup, the complement and coagulation pathways were still enriched, demonstrating that on the pathway level the alterations in the plasma proteomic profiles between the co-twins are similar between the subgroups (Figure S2). This may be an

indication that aberrant regulation of these pathways is a very early event in obesity that is independent of liver fat and perhaps precedes the accumulation of fat into the liver.

Interconnections between the significantly altered proteins and correlations between proteins in the top proteome pathways and the clinical characteristics

Because of the similarities in the pathways in both subgroups, we focused further analyses to the entire study group of 26 twin pairs. To visualize the interactions and relationships of proteins we synthesized a network in STRING that was built from 75 significantly altered proteins (Figure 3). The proteins in the top 2 STRING pathways, complement activation and regulation of inflammatory response, were highly interconnected, indicating a coordinated regulation of these pathways as a result of acquired excess body weight.

We then selected all complement and inflammatory related proteins (based on their molecular and biological functions derived from Uniprot, Table 2) amongst the 75 significantly altered proteins (Table 2) for further investigation on how they correlate with clinical adiposity and metabolic parameters on the individual level (Figure 4). Most of the significantly altered complement proteins correlated significantly positively with BMI (6/10 proteins), subcutaneous fat (5/10), intra-abdominal fat (5/10), and liver fat (5/10), and negatively with HDL (8/10). In line with previous studies, complement factor H (CFH) levels correlated significantly with insulin resistance^[37]^[44]^[45]. Further, both C3 and CFH correlated positively with liver fat. Interestingly, CFHR5 exhibited equally strong positive correlations with BMI and adiposity measurements and LDL, and a negative correlation with HDL cholesterol. CFHR5 was one the five proteins differentially expressed between the co-twins in all three study groups (all BMI-discordant as well as both the liver fat subgroups), which together with the correlation analysis, supports CFHR5 as a prospective biomarker for obesity with potential connections to dyslipidemia. In addition, CFHR5 inactivates C3b via binding to CRP^[46].

The correlations of the complement cascade proteins with the clinical measurements were overall stronger than for the inflammatory response related proteins (Figure 4). Most complement proteins correlated positively with adiposity and insulin resistance, and negatively with HDL cholesterol – similar to previously published findings^[33]. Of the inflammatory-related proteins, S100-A9 correlated positively with adiposity measures, while lipopolysaccharide-binding protein (LPS) correlated positively with adiposity measurements and negatively with insulin sensitivity, corroborating previous reports^[47]^[48]^[49]. OTU domain-binding protein 7B, an activator of the growth-promoting mTORC2 pathway^[50], previously found to be a marker candidate for early myocardial injury in plasma^[51] and inhibitor of the non-canonical NF-κB pathway^[52], correlated

positively with adiposity measurements, and negatively with HDL. HLA class II histocompatibility antigen also correlated positively with liver fat and LDL. In line with this, we previously reported in another set of obesity-discordant MZ twin pairs a significant overexpression of major histocompatibility class II molecules in adipose tissue of the heavier co-twins^[53], suggesting a role of adipocytes as antigen-presenting cells in the generation of the adaptive pro-inflammatory immune response in obesity^[54]. CFH from the complement cascade proteins and serum amyloid A1 from the inflammatory response proteins showed the strongest positive correlation with hsCRP.

Putative effects of the significantly altered complement proteins between the co-twins for the complement cascade activity

Taken together with the fact that the complement activation was the most robustly identified pathway in our analyses and clearly correlated with measurements of metabolic health, we next aimed to derive insight into what functional effects the significantly altered proteins might reflect by mapping their positions and inspecting their functions within the complement cascade (Figure 5).

The complement system is an integral part of the immune system that is increasingly recognized as a pathogenic mechanism in metabolically important tissues including adipose tissue, liver and pancreas^[55]. Complement activation can be divided into stages: Initiation through one of the three pathways (the classical, mannose-binding lectin or alternative pathway) leads to the formation of C3 convertases (Figure 5). At this stage, the initiation from any of the three pathways can be escalated through the amplification loop of the alternative pathway. Active C3 convertase results in the formation of C5 convertases, which propagate inflammation and initiate the terminal pathway culminating in the formation of membrane attack complex and cell lysis. A number of factors can inhibit the complement system to protect host tissues from complement driven damage and inflammation. In the heavier co-twins, the plasma levels of the cascade initiating proteins, and a subunit of the C3 and C5 convertases of the classical and mannose-binding lectin pathways, were increased. Additionally, the level of C3, was increased (Figure 5). The alternative pathway and the amplification loop appeared to be actively suppressed through increased levels of inhibitory proteins (CFH, CFI). Components of the terminal complement pathway (C8 and vitronectin) responsible for forming the membrane attack complex were also increased.

These findings are well in line with previous studies of serum and plasma showing obesity-related upregulation of the complement cascade^{[16] [18] [56] [57]}, with sexual dimorphism^{[17] [58]}, and in part reminiscent of our own previous study focusing on the complement system in the adipose

tissue in the same cohort of BMI-discordant monozygotic twins ^[33]. A similar pattern of increase in the classical initiation pathway (C1q), C3 and active suppression of the alternative pathway (CFH, CFI) is seen between the plasma proteome and adipose tissue transcriptome in the heavier co-twins (Figure 5, ^[33]). This observation suggests that these proteins in the plasma could be reflective of the adipose tissue status, although we acknowledge that our study set up does not allow for direct assessment of the origin of the circulating proteins. In contrast, the terminal pathway was either unaltered (C7-9) or downregulated (C6) in the adipose tissue ^[33], while the protein levels of two C8 subunits were increased in the plasma, indicative that these may reflect complement-related processes in other tissues.

Interestingly, the protein levels of plasma protease C1 inhibitor (an inhibitor of the classical initiation pathway) and vitronectin (an inhibitor of the terminal pathway ^[59]) were also increased in the heavier co-twins' plasma. Combined with the active inhibition of the amplification loop demonstrated by us and others ^[43], the increase in these inhibitory regulators could be suggestive of a wider protective mechanism acting to protect host cells and tissues from the aberrantly activated complement system, over inflammation and cell lysis.

Concluding remarks

In the present study we used a unique study set-up of BMI-discordant MZ twins to profile the effects of excess body weight on the plasma proteome. We demonstrated that obesity is associated with increased levels of proteins in the complement system and inflammatory response pathways, even in the absence of interference from genetic differences. Overall, our findings are in line with previous studies, especially the ones where plasma proteomics profiles in obesity were shown to be characterized by complement activation ^{[15] [16] [17]}. However, as our results are derived from the MZ co-twin control design, we were able to exclude the effects of heritable factors as determinants of the large interindividual variation of the plasma proteome ^[60], and of proteins in the immune ^[12] or the complement systems ^[61]. Our study highlights that complement and inflammatory proteins associate significantly with higher levels of adiposity, insulin resistance, LDL cholesterol and hsCRP as well as with lower levels of HDL cholesterol. However, although these proteomics profiles were generally following increases in all adiposity measures, they were not specific for liver fat.

The main limitation of our study was the small study sample (due to the extreme rarity of young adult healthy BMI-discordant MZ twin pairs), which resulted in a limited statistical power to identify proteins significantly differentially expressed proteins with adjusted p value or proteins specific to the accumulation of fat into the liver. While proteins with nominal significance were

included in the pathway analyses, we used three independent pathway analyzers as a way to verify the results, all of which consistently showed the complement cascade as a significantly altered pathway in obesity. The similarity of the pathway level findings in the liver fat subgroups was useful to demonstrate that the alterations in the complement pathway appear to be a very early phenomenon that precedes the detrimental ectopic fat deposition in the liver. Identifying these early events in the pathophysiology of obesity is integral to further our understanding of how healthy obesity progresses to unhealthy obesity and metabolic disease. Moreover, it should be noted that even the obese co-twins (with the exception of the one co-twin with type 2 diabetes) are considered metabolically healthy, and therefore assessments of the entire cohort are still highly informative with regard to the early molecular level changes associated with excessive body weight.

In conclusion, our findings suggest that complement dysregulation is a very early event in obesity. Furthermore, the increase in the inhibitory proteins acting on several different stages of the complement cascade could represent an endogenous protective adaptation against complement-driven inflammation and tissue damage. Thus, bolstering the inhibition of complement activation is a promising treatment target to protect against metabolic complications and preserve metabolic health in obesity.

Acknowledgements

The study was supported by grants from Academy of Finland, Centre of Excellence in Research on Mitochondria, Metabolism and Disease (grant number: 272376), Finnish Medical Foundation, Gyllenberg Foundation, Finnish Foundation for Cardiovascular Research, Novo Nordisk Foundation, University of Helsinki; Helsinki University Hospital, Finnish Diabetes Research Foundation, Academy of Finland, (grant numbers: 314383, 266286) received by KHP, Academy of Finland, Centre of Excellence in Complex Disease Genetics, (grant numbers: 129680 and 213506), National Institute of Alcohol Abuse and Alcoholism, (grant numbers: AA-09203 and AA-12502), and Academy of Finland, (grant numbers: 141054, 263278, 264146, and 265240) received by JK.

The authors have declared no conflict of interest.

Table 1. The clinical characteristics of all BMI-discordant monozygotic twin pairs (n=26) and subdivisions in the liver fat discordant (n=12) and concordant (n=14) twin pairs. Data presented as mean (SD).

	All BMI-discordant pairs			BMI-discordant, liver fat discordant pairs			BMI-discordant, liver fat concordant pairs			Comparisons between liver fat discordant vs. liver fat concordant groups		
Number of twin pairs	26			12			14			p value (Leaner vs Leaner) ^{a,b}	p value (Heavier vs Heavier) ^{a, b}	p value (Within-pair differences) ^{a,b}
Age range	22.8-37.0			22.8-36.1			24.1-37			0.8679	0.8679	
Sex (m/f)	9/17			6/6			3/11			0.127	0.127	
Clinical Parameters	Leaner	Heavier	Within-Pair	Leaner	Heavier	Within-Pair	Leaner	Heavier	Within-Pair			
			p value (Leaner vs Heavier) ^c			p value (Leaner vs Heavier) ^c			p value (Leaner vs Heavier) ^c			
Height (cm)	172.5 (11.1)	172.8 (10.8)	0.2827	173.1 (11.3)	173.4 (11.1)	0.5004	171.9 (11.2)	172.3 (11.0)	0.4265	0.7850	0.7936	0.9066
Weight (kg)	74.9 (17.9)	92.6 (20.0)	<0.001	77.4 (18.7)	98.1 (20.9)	<0.001	72.7 (17.6)	87.9 (18.7)	<0.001	0.5051	0.2038	0.1254
BMI (kg/m ²)	24.9 (4.4)	30.8 (5.1)	<0.001	25.6 (3.7)	32.4 (4.8)	<0.001	24.4 (5.0)	29.4 (5.1)	<0.001	0.5147	0.1411	0.0930
Body fat (%)	31.1 (10.1)	40.7 (7.2)	<0.001	29.0 (11.2)	39.0 (7.2)	<0.001	32.9 (9.1)	41.9 (7.2)	<0.001	0.3415	0.3444	0.6027
Subcutaneous adipose tissue (cm ³)	3587.1 (2120.5)	6102.6 (2758.7)	<0.001	3513.2 (2209.5)	6240.4 (3332.49)	<0.001	3650.5 (2122.9)	5984.6 (2281.9)	<0.001	0.8731	0.8191	0.5398
Intra-abdominal adipose tissue (cm ³)	591.6 (396.8)	1387.1 (770.9)	<0.001	623.8 (350.1)	1874.5 (763.1)	<0.001	564.0 (444.1)	969.2 (492.6)	<0.001	0.7095	0.0013	<0.001
Liver fat (%)	0.8 (0.9)	3.9 (4.0)	0.0006	1.0 (0.9)	7.3 (3.3)	<0.001	0.7 (0.8)	0.9 (0.7)	0.1570	0.4731	<0.001	<0.001
fP-glucose	5.1 (0.4)	5.4 (1.0)	0.1182	5.2 (0.5)	5.7 (1.4)	0.1590	5.0 (0.3)	5.1 (0.5)	0.5448	0.3224	0.1346	0.2211

Received: 10 18, 2018; Revised: 12 27, 2018; Accepted: 01 24, 2019

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1002/prca.201800173](#).

This article is protected by copyright. All rights reserved.

(mmol/L)												
fS-insulin (mU/L)	5.1 (2.9)	7.9 (4.5)	0.0027	5.1 (2.7)	9.6 (5.2)	0.0039	5.1 (3.3)	6.5 (3.4)	0.2156	0.9745	0.0820	0.0588
HOMA index	1.1 (0.6)	1.9 (1.2)	0.0003	1.2 (0.7)	2.5 (1.4)	0.0031	1.0 (0.6)	1.5 (0.9)	0.0394	0.4455	0.0411	0.0568
Matsuda index	9.6 (4.9)	5.8 (3.1)	0.0044	9.6 (5.6)	4.2 (1.3)	0.0031	9.6 (4.4)	7.1 (3.6)	0.1745	0.9555	0.0203	0.2410
Total cholesterol (mmol/L)	4.3 (0.7)	4.6 (0.9)	0.1334	4.2 (0.5)	5.0 (0.8)	0.0273	4.4 (0.8)	4.4 (0.9)	0.8986	0.6084	0.0763	0.0618
LDL cholesterol (mmol/L)	2.5 (0.6)	3.0 (0.8)	0.0136	2.3 (0.4)	3.2 (0.9)	0.0075	2.7 (0.7)	2.8 (0.7)	0.5758	0.1416	0.1821	0.0248
HDL cholesterol (mmol/L)	1.6 (0.4)	1.3 (0.4)	0.0002	1.7 (0.4)	1.3 (0.5)	0.0030	1.6 (0.3)	1.3 (0.4)	0.0232	0.4724	0.8468	0.2902
Triglycerides (mmol/L)	0.9 (0.3)	1.2 (0.5)	0.0165	0.8 (0.3)	1.4 (0.6)	0.0304	0.9 (0.3)	1.0 (0.3)	0.3033	0.5838	0.0673	0.0689
High- sensitivity CRP (mg/L)	1.8 (2.5)	3.6 (5.7)	0.0580	1.9 (2.6)	5.2 (7.3)	0.0381	1.8 (2.5)	1.9 (2.4)	0.9346	0.9721	0.2152	0.0639
Physical activity index	8.6 (1.5)	7.9 (1.4)	0.0119	8.8 (1.6)	7.6 (1.1)	0.0212	8.5 (1.4)	8.1 (1.6)	0.2350	0.5609	0.4455	0.2254
Alcohol intake (dose/week)	5.4 (10.0)	9.0 (8.8)	0.0961	3.8 (2.8)	11.0 (9.8)	0.0209	7.2 (14.4)	6.7 (7.5)	0.8631	0.4754	0.2972	0.0557
Energy intake (kcal/day)	2126.3 (540.2)	2181.1 (577.1)	0.6900	2070.3 (643.4)	2258.4 (540.1)	0.4386	2170.3 (464.3)	2120.3 (617.6)	0.7610	0.5825	0.5638	0.8549

^a Unpaired t-tests were used for computing comparisons between liver fat discordant and concordant pairs (leaner vs leaner, heavier vs heavier and of within-pair differences).

^b Chi2 test was used to test the sex difference between liver fat discordant vs. concordant groups.

^c *Paired t-tests were used for computing comparisons of leaner and heavier co-twins within twin pairs.*

Table 2. The 75 significantly altered proteins in the plasma proteomics between leaner and heavier BMI-discordant monozygotic co-twins (n= 26 twin pairs).

Data derived from Uniprot. Proteins that are selected for correlation analysis are indicated in bold.

Gene name	Accession	Protein description	Molecular function	Biological process	Log Fold Change	p value	adj p value
KRT87P	A6NCN2	Putative keratin-87 protein	structural molecule activity		0.51	0.041	0.28
HLA-DOA	P06340	HLA class II histocompatibility antigen	MHC class II receptor activity	Immune response	0.34	0.006	0.10
DLG2	Q15700	Disks large homolog 2	kinase binding		0.33	0.004	0.10
GOLGA3	Q08378	Golgin subfamily A member 3	transporter activity	transport, cellular localization	0.33	0.024	0.19
EVL	Q9UI08	Ena/VASP-like protein	actin binding	cytoskeleton remodeling	0.32	0.012	0.14
PROZ	P22891	Vitamin K-dependent protein Z	receptor binding	Blood coagulation, Hemostasis	0.3	0.008	0.12
RAE1	P78406	mRNA export factor	RNA binding	cell cycle	0.3	0.013	0.14
EIF3H	O15372	Eukaryotic translation initiation factor 3 subunit H	RNA binding	translational initiation, protein synthesis	0.29	0.022	0.17
F7	P08709	Coagulation factor VII	receptor binding	complement and coagulation cascades	0.28	0.002	0.09
MKS1	Q9NXB0	Meckel syndrome type 1 protein		Cilium biogenesis/degradation	0.26	0.016	0.16
ANKIB1	Q9P2G1	Ankyrin repeat and IBR domain-containing protein 1	Transferase	protein polyubiquitination	0.2	0.012	0.14
PCYOX1	Q9UHG3	Prenylcysteine oxidase 1	Oxidoreductase		0.2	0.013	0.14
TUBB4A	P04350	Tubulin beta-4A chain	structural constituent of cytoskeleton	microtubule-based process	0.19	0.046	0.29
PTGDS	P41222	Prostaglandin-H2 D-isomerase	fatty acid binding	Fatty acid biosynthesis, Fatty acid metabolism	0.17	0.020	0.17

OTUD7B	Q6GQQ9	OTU domain-containing protein 7B	ubiquitin binding	immune response, NF-kappaB signaling	0.17	0.004	0.10
LBP	P18428	Lipopolysaccharide-binding protein	lipopeptide binding	immune response,inflammatory response	0.16	0.001	0.09
VPS11	Q9H270	Vacuolar protein sorting-associated protein 11 homolog	protein binding	lysosome organization,protein trafficking	0.15	0.009	0.13
HGFAC	Q04756	Hepatocyte growth factor activator	serine-type endopeptidase activity	proteolysis	0.15	0.003	0.10
SAA1	P0DJI8	Serum amyloid A-1 protein	Heparin-binding	inflammatory response	0.15	0.020	0.17
BLVRB	P30043	Flavin reductase (NADPH)	Oxidoreductase	small molecule metabolic process	0.14	0.005	0.10
NUP107	P57740	Nuclear pore complex protein	RNA binding	RNA transport, mRNA transport	0.13	0.0002	0.08
C4B	P0C0L5	Complement C4-B	complement binding	complement activation	0.12	0.001	0.08
CFHR5	Q9BXR6	Complement factor H-related protein 5		complement activation	0.12	0.002	0.10
WDR19	Q8NEZ3	WD repeat-containing protein 19		Cilium biogenesis/degradation	0.12	0.001	0.08
DOCK9	Q9BZ29	Dedicator of cytokinesis protein 9	cadherin binding	blood coagulation	0.12	0.017	0.16
PIP4K2C	Q8TBX8	Phosphatidylinositol 5-phosphate 4-kinase type-2 gamma	Kinase, Transferase	inositol phosphate metabolic process, lipid metabolic process	0.12	0.048	0.30
TXNRD2	Q9NNW7	Thioredoxin reductase 2, mitochondrial	Oxidoreductase	cell redox homeostasis	0.12	0.044	0.29
FBXO22	Q8NEZ5	F-box only protein 22	ubiquitin-protein transferase activity	cellular protein catabolic process	0.12	0.001	0.09
TINAG	Q9UJW2	Tubulointerstitial nephritis antigen	receptor binding	Cell adhesion	0.12	0.015	0.15
KRT82	Q9NSB4	Keratin, type II		keratinization	0.11	0.007	0.12
MED16	Q9Y2X0	Mediator of RNA polymerase II transcription subunit 16	catalytic activity	Transcription regulation	0.11	0.003	0.10
LDHA	P00338	L-lactate dehydrogenase A chain	Oxidoreductase	oxidation-reduction process	0.1	0.047	0.30
GTFC4	Q9UKN8	General transcription factor 3C polypeptide 4	Transferase	transcription	0.1	0.003	0.10
HEXB	P07686	Beta-hexosaminidase subunit beta	Glycosidase, Hydrolase	Glycosaminoglycan degradation	0.1	0.006	0.10

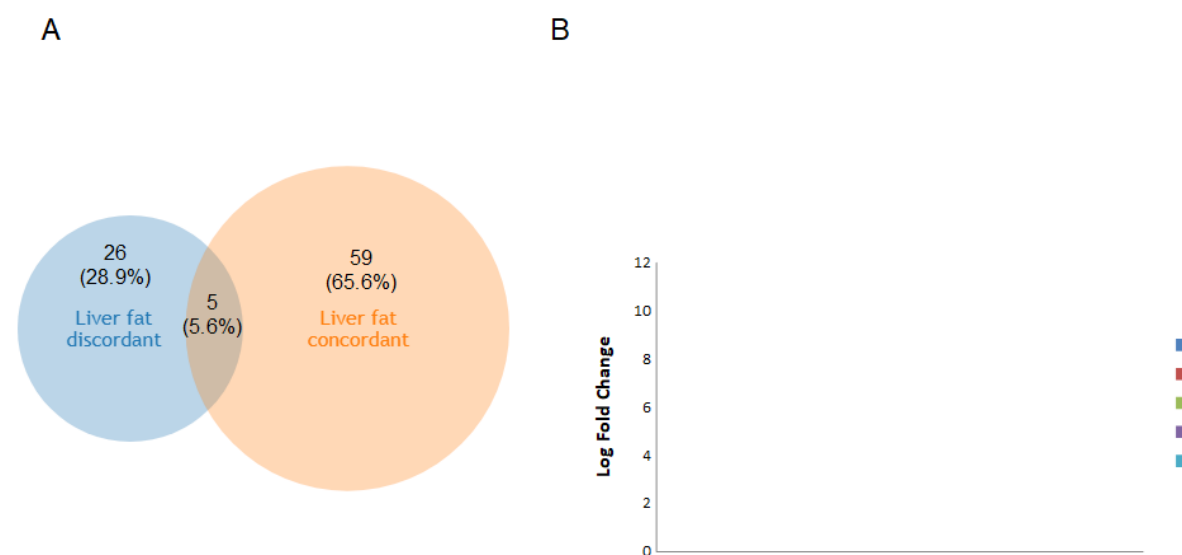
SERPIND1	P05546	Heparin cofactor 2	heparin binding	blood coagulation, complement and coagulation cascades	0.1	0.003	0.10
BCHE	P06276	Cholinesterase	choline binding	choline metabolic process	0.1	0.017	0.16
ADAMTS12	P58397	A disintegrin and metalloproteinase with thrombospondin motifs 12	metalloendopeptidase activity	cell migration, cell response to insulin	0.09	0.003	0.10
C1QA	P02745	Complement C1q subcomponent subunit A	complement component C1q binding	complement activation	0.09	0.019	0.17
TTK	P33981	Dual specificity protein kinase TTK	Tyrosine-protein kinase	cell division	0.09	0.048	0.30
THBS1	P07996	Thrombospondin-1	receptor binding	cell adhesion, platelet degranulation	0.09	0.009	0.12
SERPING1	P05155	Plasma protease C1 inhibitor	Protease inhibitor	complement activation	0.09	0.026	0.20
MASP1	P48740	Mannan-binding lectin serine protease 1	Hydrolase	complement activation	0.09	0.010	0.13
PTGR2	Q8N8N7	Prostaglandin reductase 2	catalytic activity	prostaglandin metabolic process	0.09	0.043	0.29
APOB	P04114	Apolipoprotein B-100	cholesterol transporter activity	Regulation of plasma, lipoprotein level, lipid homeostasis	0.09	0.001	0.08
CFH	P08603	Complement factor H		complement activation	0.08	0.008	0.12
CFI	P05156	Complement factor I	complement binding	complement activation	0.08	0.005	0.10
IZUMO3	Q5VZ72	Izumo sperm-egg fusion protein 3	protein homodimerization activity		0.08	0.005	0.10
S100A9	P06702	Protein S100-A9	receptor binding	inflammatory response	0.08	0.011	0.14
C3	P01024	Complement C3	receptor binding	complement activation	0.07	0.002	0.09
NUP210L	Q5VU65	Nuclear pore membrane glycoprotein 210-like		cell development	0.07	0.049	0.30
C8B	P07358	Complement component C8 beta chain	protein complex binding	complement activation	0.06	0.033	0.23
THOC3	Q96J01	THO complex subunit 3	RNA binding	RNA transport	0.06	0.045	0.29
C8G	P07360	Complement component C8 gamma chain	complement binding	complement activation	0.05	0.045	0.29
POTEI	P0CG38	POTE ankyrin domain family member I	ion binding	retina homeostasis	0.05	0.022	0.17

VTN	P04004	Vitronectin	Heparin-binding	Cell adhesion	0.05	0.031	0.23
SERPINF1	P36955	Pigment epithelium-derived factor	serine-type endopeptidase inhibitor activity	cellular component organization	0.05	0.015	0.15
APOD	P05090	Apolipoprotein D	cholesterol binding	lipid transport	-0.07	0.011	0.14
LIPH	Q8WWY8	Lipase member H	phospholipase activity	lipid catabolic process	-0.11	0.004	0.10
PLGLA	Q15195	Plasminogen-like protein A	Protein binding		-0.12	0.019	0.16
CCDC102B	Q68D86	Coiled-coil domain-containing protein 102B			-0.12	0.018	0.16
LRR9	Q6ZRR7	Leucine-rich repeat-containing protein 9		multicellular organismal homeostasis	-0.13	0.033	0.23
SHBG	P04278	Sex hormone-binding globulin	Hormone binding	Regulates the plasma	-0.17	0.018	0.16
ZNF778	P60409	Zinc finger protein 778	DNA binding, Metal ion binding	Transcription regulation	-0.19	0.006	0.10
RNF112	Q9ULX5	RING finger protein 112	GTPase activity	Cell cycle	-0.19	0.026	0.20
GAK	O14976	Cyclin-G-associated kinase	Nucleotide-binding, Kinase activity	Cell cycle	-0.19	0.027	0.20
IGJ	P01591	Immunoglobulin J chain	receptor binding	receptor binding	-0.2	0.011	0.14
ANAPC4	Q9UJX5	Anaphase-promoting complex subunit 4	protein phosphatase binding	cell cycle	-0.27	0.008	0.12
PKM	P14618	Pyruvate kinase PKM	kinase activity	Metabolic pathways	-0.27	0.017	0.16
DNASE1L1	P49184	Deoxyribonuclease-1-like 1	DNA binding	DNA metabolic process, DNA binding	-0.27	0.009	0.12
CROCC	Q5TZA2	Rootletin	Actin binding, Protein binding	Centrosome cycle, cell cycle, Cilium biogenesis	-0.28	0.012	0.14
IFT43	Q96FT9	Intraflagellar transport protein 43		Cilium biogenesis	-0.31	0.006	0.10
UFL1	O94874	E3 UFM1-protein ligase 1	ligase activity	Ubiquitin mediated proteolysis	-0.34	0.004	0.10
IGKC	P01834	Immunoglobulin kappa constant	antigen binding	complement activation	-0.36	0.050	0.30
SH2D3A	Q9BRG2	SH2 domain-containing protein 3A	guanyl-nucleotide exchange factor activity	JNK cascade, protein targeting to nucleus	-0.43	0.004	0.10
GALNT6	Q8NCL4	Polypeptide N-acetylgalactosaminyl	N-	protein glycosylation	-0.46	0.030	0.22

		transferase 6	acetylgalactosaminyltransferase activity				
--	--	---------------	---	--	--	--	--

Figures legends

Figure 1. A) *Venn diagram of significantly altered proteins in liver fat discordant and concordant monozygotic twins' plasma. *[Oliveros, J.C. (2007) VENNY. An interactive tool for comparing lists with Venn Diagrams. <http://bioinfogp.cnb.csic.es/tools/venny/index.html>]. B) Logarithmic fold-changes (heavier co-twins vs leaner co-twins) for the 5 common proteins in BMI-discordant, liver fat discordant and liver fat concordant subgroups. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. APOB: apolipoprotein B100, C4B: complement c4b, CFHR5: complement factor H-related protein 5, F7: coagulation factor VII, NUP 107: nuclear pore complex protein Nup107.



Received: 10 18, 2018; Revised: 12 27, 2018; Accepted: 01 24, 2019

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1002/prca.201800173](https://doi.org/10.1002/prca.201800173).

This article is protected by copyright. All rights reserved.

Figure 2. Pathway analyses of the 75 significantly altered proteins in heavy vs lean comparison of all BMI-discordant monozygotic twins' plasma (n=26 twin pairs). A) Top 5 significantly enriched biological process according to DAVID. B) Top 5 significantly enriched canonical pathway according to IPA. C) Top 5 significantly enriched biological processes according to STRING.

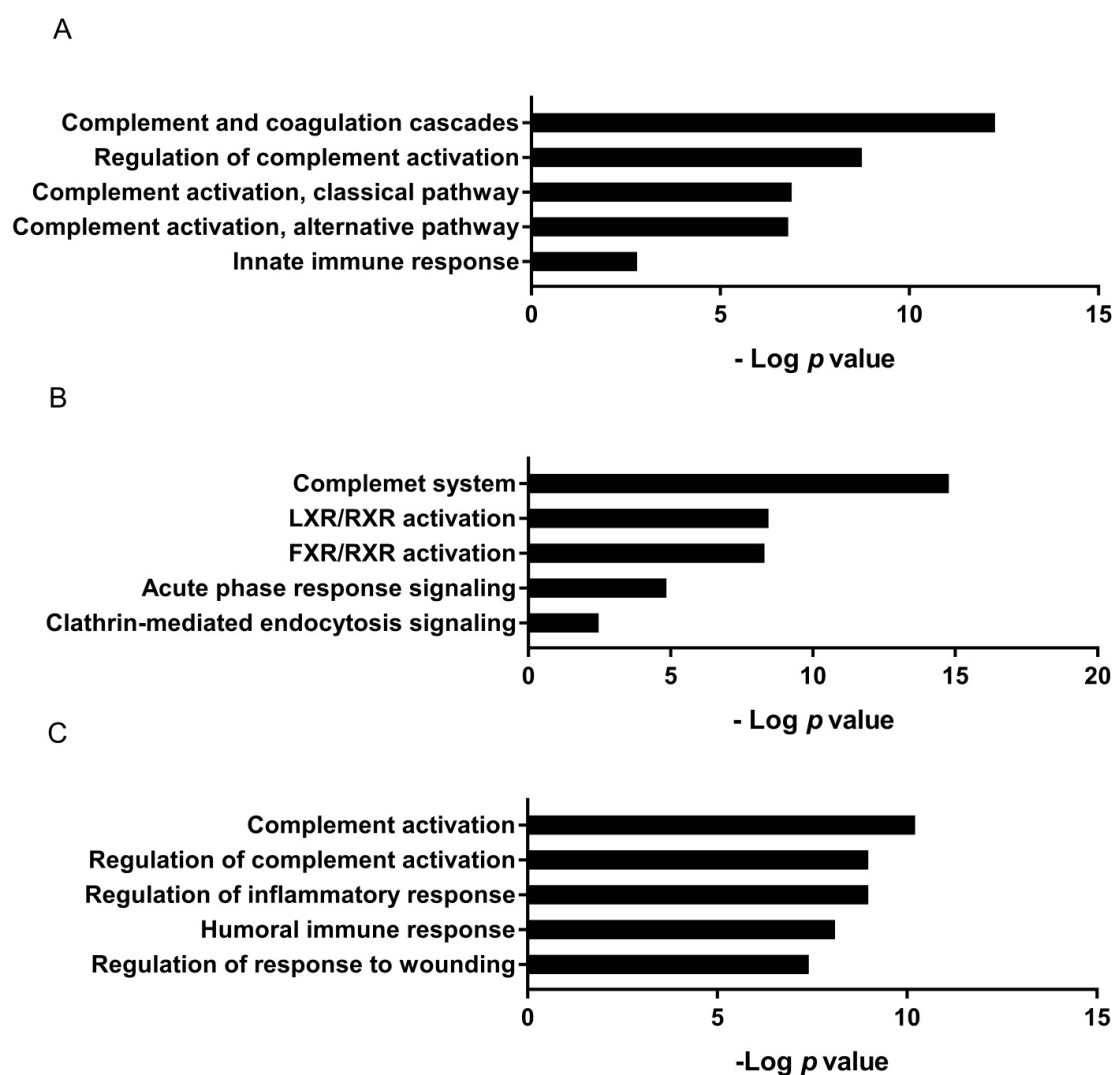


Figure 3. Network of 75 significantly altered proteins in lean and heavy BMI-discordant monozygotic twins' plasma (n=26 twin pairs). Top scoring pathways derived from STRING are highlighted: complement activation (red nodes) and regulation of inflammatory response (blue nodes). Local coefficient clustering: 0.411, PPI enrichment p value: 7.86×10^{-9} , expected number of edges: 29, number of edges: 64 (this indicates that our proteins have more interactions among themselves than what would be expected for a random set of proteins of similar size). The thickness of the line between nodes indicates the strength of data support. Disconnected nodes are removed. Data derived from STRING ver. 10.5 (<http://string-db.org>).

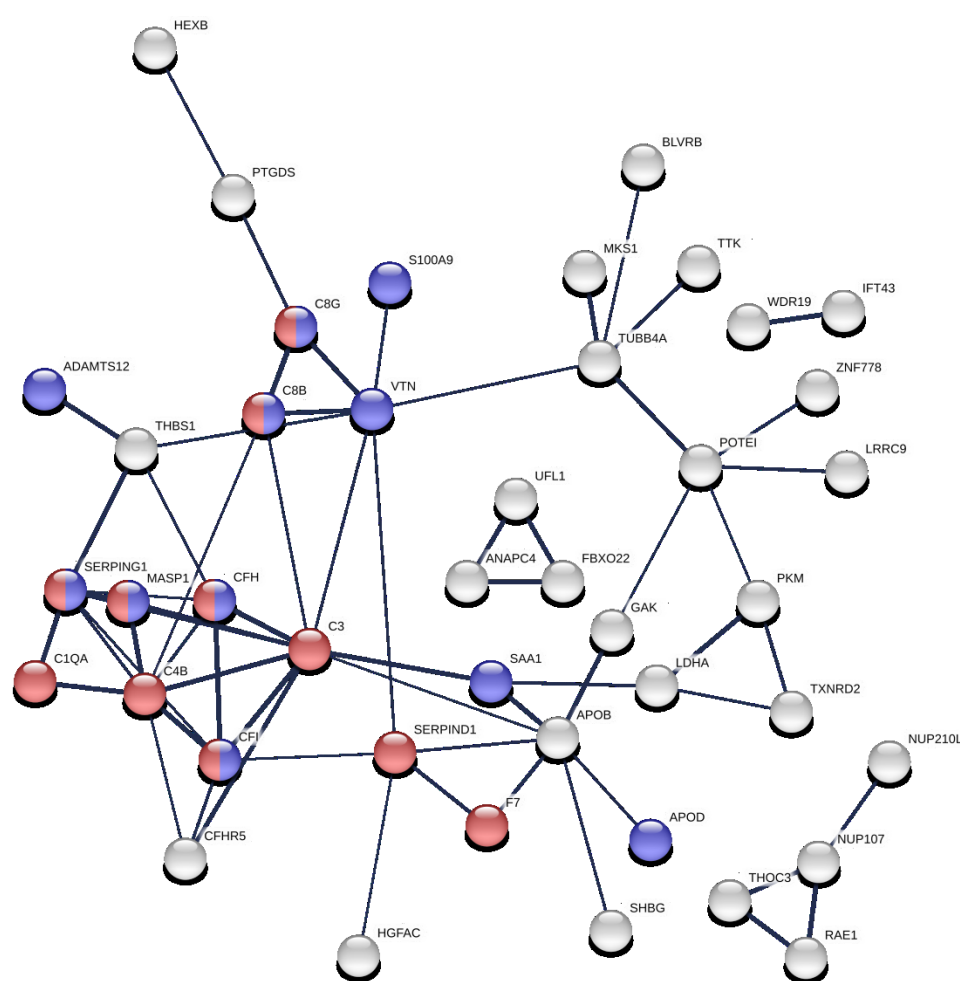


Figure 4. The correlations of the complement and inflammatory response related proteins with the adiposity and metabolic measurements in monozygotic twin individuals (n= 52). Proteins are selected due to their molecular function derived from Uniprot. BMI: Body mass index, Sc fat: subcutaneous fat, Ia fat: intra-abdominal fat, HOMA: The Homeostasis Model Assessment, LDL: Low density lipoprotein, HDL: High density lipoprotein. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (after multiple correction).

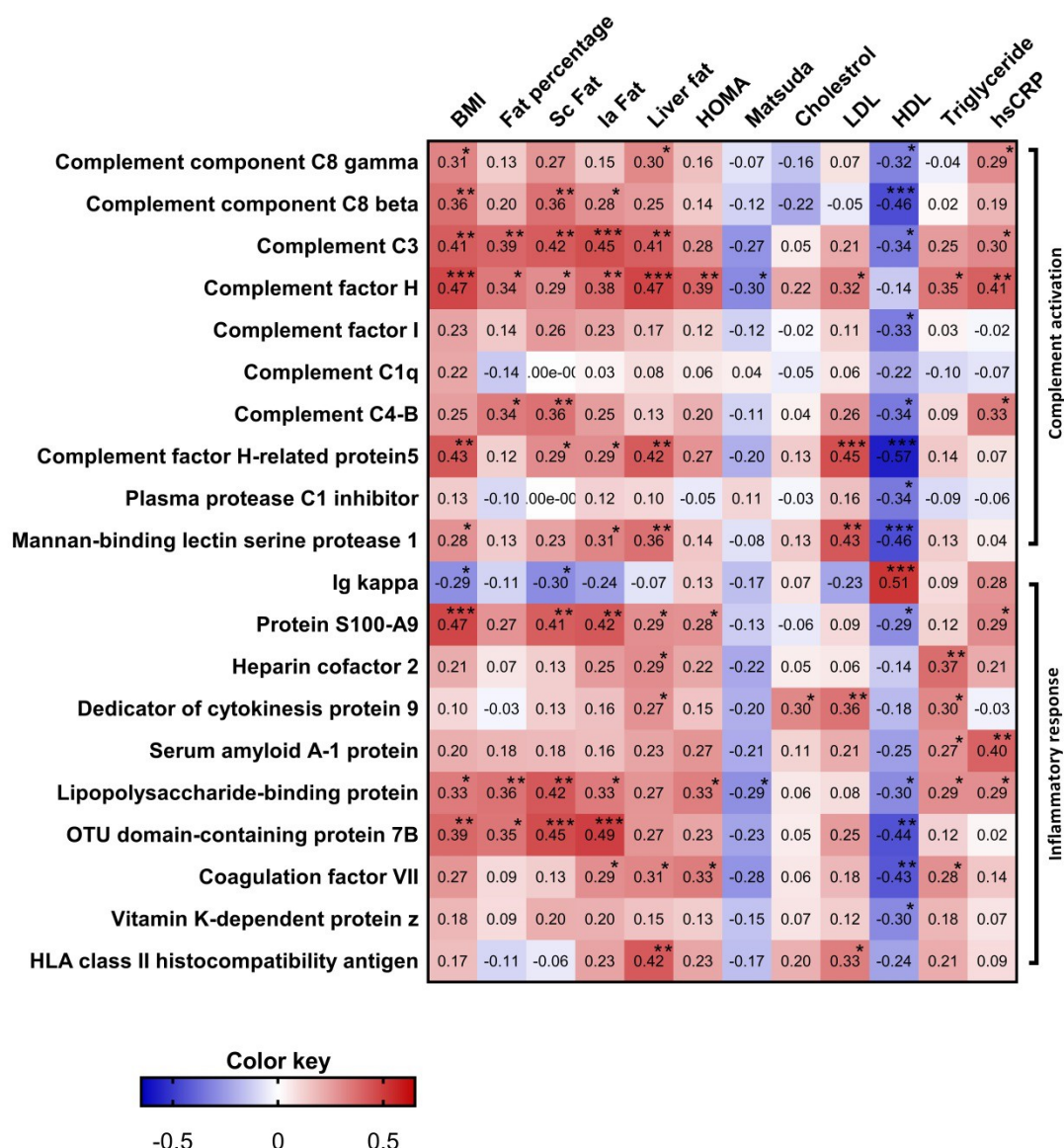
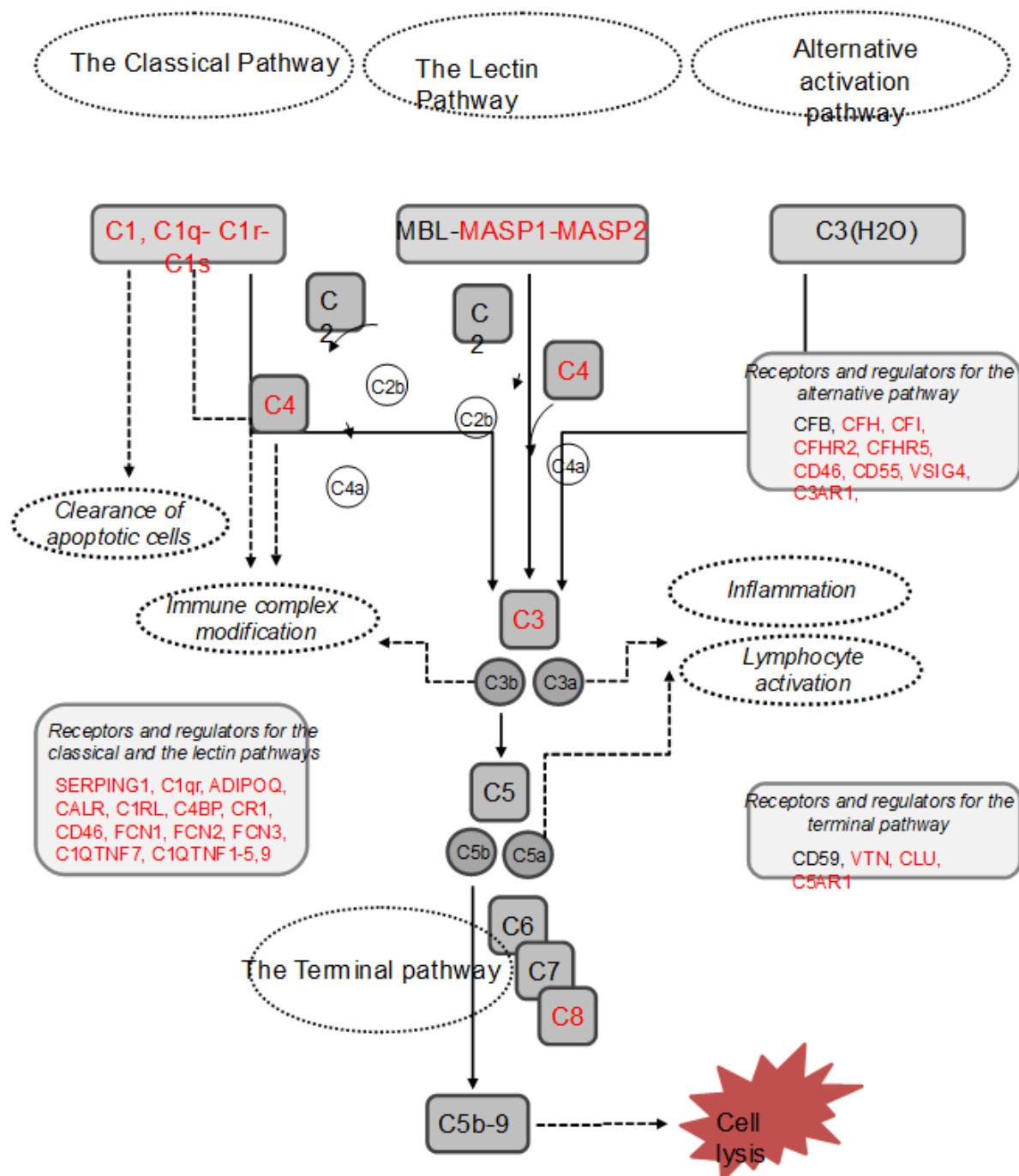


Figure 5. The complement cascade protein profile in plasma in monozygotic twin individuals (n= 52). Proteins that are upregulated in heavier co-twins are indicated with red.



Supplementary Figures and tables legends

Figure S1. Liver fat content measured by magnetic resonance spectroscopy in 26 monozygotic BMI-discordant pairs (BMI difference > 3 kg/m²). 12 pairs were grouped as liver fat discordant and 14 pairs as liver fat concordant.

Figure S2. Biological pathway analysis of significantly expressed proteins in the lean vs. heavy co-twin comparison of liver fat discordant (n = 12 twin pairs) and liver fat concordant (n = 14 twin pairs) BMI-discordant monozygotic twin pair subgroups. Top significantly enriched biological process based on STRING, DAVID and IPA analysis.

Table S1-S4. Differentially expressed plasma proteins between the lean and heavy co-twins in liver fat discordant (n=12 twin pairs) and liver fat concordant (n=14 twin pairs) monozygotic twin pair subgroups.

References

- [1] M. Ng, T. Fleming, M. Robinson, Lancet 2014, 384, 746.
- [2] A. J. Klil-Drori, L. Azoulay, M. N. Pollak, Nat Rev Clin Oncol 2017, 14, 85.
- [3] V. Primeau, L. Coderre, A. D. Karelis, M. Brochu, M. E. Lavoie, V. Messier, R. Sladek, R. Rabasa-Lhoret, Int J Obes (Lond) 2011, 35, 971.
- [4] H. Rudolph, S. Bluher, C. Falkenberg, M. Neef, A. Korner, J. Wurz, W. Kiess, E. Brahler, Obes Facts 2010, 3, 83.
- [5] G. Bedogni, L. Miglioli, F. Masutti, C. Tiribelli, G. Marchesini, S. Bellentani, Hepatology 2005, 42, 44.
- [6] N. Stefan, K. Kantartzis, J. Machann, F. Schick, C. Thamer, K. Rittig, B. Balletshofer, F. Machicao, A. Fritsche, H. U. Haring, Arch Intern Med 2008, 168, 1609.
- [7] A. Kotronen, H. Yki-Jarvinen, K. Sevastianova, R. Bergholm, A. Hakkarainen, K. H. Pietilainen, L. Juurinen, N. Lundbom, T. I. Sorensen, Obesity (Silver Spring) 2011, 19, 23.
- [8] J. Naukkarinen, S. Heinonen, A. Hakkarainen, J. Lundbom, K. Vuolteenaho, L. Saarinen, S. Hautaniemi, A. Rodriguez, G. Fruhbeck, P. Pajunen, T. Hyotylainen, M. Oresic, E. Moilanen, A. Suomalainen, N. Lundbom, J. Kaprio, A. Rissanen, K. H. Pietilainen, Diabetologia 2014, 57, 167.
- [9] M. Mann, N. A. Kulak, N. Nagaraj, J. Cox, Mol Cell 2013, 49, 583.
- [10] R. Aebersold, M. Mann, Nature 2016, 537, 347.
- [11] H. Zhang, A. Y. Liu, P. Loriaux, B. Wollscheid, Y. Zhou, J. D. Watts, R. Aebersold, Mol Cell Proteomics 2007, 6, 64.
- [12] Y. S. Liu, A. Buil, B. C. Collins, L. C. J. Gillet, L. C. Blum, L. Y. Cheng, O. Vitek, J. Mouritsen, G. Lachance, T. D. Spector, E. T. Dermitzakis, R. Aebersold, Mol Syst Biol 2015, 11.
- [13] C. B. Garrison, K. J. Lastwika, Y. Z. Zhang, C. I. Li, P. D. Lampe, J Proteome Res 2017, 16, 274.
- [14] P. E. Geyer, L. M. Holdt, D. Teupser, M. Mann, Mol Syst Biol 2017, 13.
- [15] M. Abu-Farha, A. Tiss, J. Abubaker, A. Khadir, F. Al-Ghimlas, I. Al-Khairi, E. Baturcam, P. Cherian, N. Elkum, M. Hammad, J. John, S. Kavalakatt, S. Warsame, K. Behbehani, S. Dermime, M. Dehbi, Plos One 2013, 8.

- [16] A. Oberbach, M. Bluher, H. Wirth, H. Till, P. Kovacs, Y. Kullnick, N. Schlichting, J. M. Tamm, U. Rolle-Kampczyk, J. Murugaiyan, H. Binder, A. Dietrich, M. von Bergen, *J Proteome Res* 2011, 10, 4769.
- [17] A. M. Curran, C. F. Draper, M. P. Scott-Boyer, A. Valsesia, H. M. Roche, M. F. Ryan, M. J. Gibney, M. Kutmon, C. T. Evelo, S. L. Coort, A. Astrup, W. H. Saris, L. Brennan, J. Kaput, *J Proteome Res* 2017, 16, 4122.
- [18] P. E. Geyer, N. J. W. Albrechtsen, S. Tyanova, N. Grassl, E. W. Iepsen, J. Lundgren, S. Madsbad, J. J. Holst, S. S. Torekov, M. Mann, *Mol Syst Biol* 2016, 12.
- [19] J. Kaprio, *Twin Res Hum Genet* 2006, 9, 772.
- [20] M. Graner, A. Seppala-Lindroos, A. Rissanen, A. Hakkarainen, N. Lundbom, J. Kaprio, M. S. Nieminen, K. H. Pietilainen, *The American journal of cardiology* 2012, 109, 1295.
- [21] M. Muniandy, S. Heinonen, H. Yki-Jarvinen, A. Hakkarainen, J. Lundbom, N. Lundbom, J. Kaprio, A. Rissanen, M. Ollikainen, K. H. Pietilainen, *Int J Obesity* 2017, 41, 1176.
- [22] A. Pietrobello, C. Formica, Z. M. Wang, S. B. Heymsfield, *Am J Physiol-Endoc M* 1996, 271, E941.
- [23] J. C. Hannukainen, R. Borra, K. Linderborg, H. Kallio, J. Kiss, V. Lepomaki, K. K. Kalliokoski, U. M. Kujala, J. Kaprio, O. J. Heinonen, M. Komu, R. Parkkola, M. Ahotupa, T. Lehtimäki, R. Huupponen, P. Iozzo, P. Nuutila, *J Hepatol* 2011, 54, 545.
- [24] J. A. H. Baecke, J. Burema, J. E. R. Frijters, *Am J Clin Nutr* 1982, 36, 936.
- [25] S. Heinonen, L. Saarinen, J. Naukkarinen, A. Rodriguez, G. Fruhbeck, A. Hakkarainen, J. Lundbom, N. Lundbom, K. Vuolteenaho, E. Moilanen, P. Arner, S. Hautaniemi, A. Suomalainen, J. Kaprio, A. Rissanen, K. H. Pietilainen, *Int J Obesity* 2014, 38, 1423.
- [26] M. Saraswat, S. Joenvaara, H. Seppanen, H. Mustonen, C. Haglund, R. Renkonen, *Cancer Med-Us* 2017, 6, 1738.
- [27] J. C. Silva, M. V. Gorenstein, G. Z. Li, J. P. C. Vissers, S. J. Geromanos, *Mol Cell Proteomics* 2006, 5, 144.
- [28] G. K. Smyth, *Stat Appl Genet Mol Biol* 2004, 3, Article3.
- [29] B. Phipson, S. Lee, I. J. Majewski, W. S. Alexander, G. K. Smyth, *Ann Appl Stat* 2016, 10, 946.
- [30] K. Kammers, R. N. Cole, C. Tiengwe, I. Ruczinski, *EuPA Open Proteom* 2015, 7, 11.
- [31] N. D. Price, A. T. Magis, J. C. Earls, G. Glusman, R. Levy, C. Lausted, D. T. McDonald, U. Kusebauch, C. L. Moss, Y. Zhou, S. Qin, R. L. Moritz, K. Brogaard, G. S. Omenn, J. C. Lovejoy, L. Hood, *Nat Biotechnol* 2017, 35, 747.
- [32] J. A. Vizcaino, A. Csordas, N. del-Toro, J. A. Dienes, J. Griss, I. Lavidas, G. Mayer, Y. Perez-Riverol, F. Reisinger, T. Ternent, Q. W. Xu, R. Wang, H. Hermjakob, *Nucleic Acids Res* 2016, 44, D447.
- [33] S. Kaye, A. I. Lokki, A. Hanttu, E. Nissila, S. Heinonen, A. Hakkarainen, J. Lundbom, N. Lundbom, L. Saarinen, O. Tynnenen, M. Muniandy, A. Rissanen, J. Kaprio, S. Meri, K. H. Pietilainen, *Front Immunol* 2017, 8.
- [34] S. Jukarainen, S. Heinonen, J. T. Ramo, R. Rinnankoski-Tuikka, E. Rappou, M. Tummers, M. Muniandy, A. Hakkarainen, J. Lundbom, N. Lundbom, J. Kaprio, A. Rissanen, E. Pirinen, K. H. Pietilainen, *J Clin Endocr Metab* 2016, 101, 274.
- [35] J. T. Ramo, S. M. Kaye, S. Jukarainen, L. H. Bogl, A. Hakkarainen, J. Lundbom, N. Lundbom, A. Rissanen, J. Kaprio, N. Matikainen, K. H. Pietilainen, *The Journal of clinical endocrinology and metabolism* 2017, 102, 220.
- [36] B. Klop, J. W. Elte, M. C. Cabezas, *Nutrients* 2013, 5, 1218.
- [37] J. M. Moreno-Navarrete, J. M. Fernandez-Real, *Semin Cell Dev Biol* 2017.
- [38] J. McEneny, J. A. Daniels, A. McGowan, A. Gunness, K. Moore, M. Stevenson, I. S. Young, J. Gibney, *J Diabetes Res* 2015.
- [39] P. M. Ridker, N. Rifai, L. Rose, J. E. Buring, N. R. Cook, *N Engl J Med* 2002, 347, 1557.
- [40] G. A. Rosito, R. B. D'Agostino, J. Massaro, L. Lipinska, M. A. Mittleman, P. Sutherland, P. W. F. Wilson, D. Levy, J. E. Muller, G. H. Tofler, *Thromb Haemostasis* 2004, 91, 683.

- [41] I. F. Godsland, D. Crook, A. J. Proudler, J. C. Stevenson, *J Clin Endocr Metab* 2005, 90, 190.
- [42] P. Kuusela, M. Saraswat, S. Joenvaara, J. Kaartinen, A. Jarvinen, R. Renkonen, *Plos One* 2017, 12, e0172987.
- [43] S. O. Moreno, O. Cominetti, A. N. Galindo, I. Irincheeva, J. Cortesy, A. Astrup, W. M. Saris, J. Hager, M. Kussmann, L. Dayon, *Proteom Clin Appl* 2018, 12.
- [44] M. M. J. van Greevenbroek, M. Jacobs, C. J. H. van der Kallen, V. M. M. J. Vermeulen, E. H. J. M. Jansen, C. G. Schalkwijk, I. Ferreira, E. J. M. Feskens, C. D. A. Stehouwer, *Eur J Clin Invest* 2011, 41, 372.
- [45] H. Esterbauer, F. Krempler, H. Oberkofler, W. Patsch, *Eur J Clin Invest* 1999, 29, 653.
- [46] C. T. Park, S. D. Wright, *J Biol Chem* 1996, 271, 18054.
- [47] O. H. Mortensen, A. R. Nielsen, C. Erikstrup, P. Plomgaard, C. P. Fischer, R. Krogh-Madsen, B. Lindegaard, A. M. Petersen, S. Taudorf, B. K. Pedersen, *Plos One* 2009, 4.
- [48] J. M. Moreno-Navarrete, F. Ortega, M. Serino, E. Luche, A. Waget, G. Pardo, J. Salvador, W. Ricart, G. Fruhbeck, R. Burcelin, J. M. Fernandez-Real, *Int J Obesity* 2012, 36, 1442.
- [49] L. A. Sun, Z. J. Yu, X. W. Ye, S. R. Zou, H. X. Li, D. X. Yu, H. Y. Wu, Y. Chen, J. Dore, K. Clement, F. B. Hu, X. Lin, *Diabetes Care* 2010, 33, 1925.
- [50] B. Wang, Z. L. Jie, D. H. Joo, A. Ordureau, P. Liu, W. J. Gan, J. P. Guo, J. F. Zhang, B. J. North, X. P. Dai, X. H. Cheng, X. W. Bian, L. Q. Zhang, J. W. Harper, S. C. Sun, W. Y. Wei, *Nature* 2017, 545, 365.
- [51] H. Keshishian, M. W. Burgess, M. A. Gillette, P. Mertins, K. R. Clauser, D. R. Mani, E. W. Kuhn, L. A. Farrell, R. E. Gerszten, S. A. Carr, *Mol Cell Proteomics* 2015, 14, 2375.
- [52] Z. J. J. Chen, *Nat Cell Biol* 2005, 7, 758.
- [53] K. H. Pietilainen, J. Naukkarinen, A. Rissanen, J. Saharinen, P. Ellonen, H. Keranen, A. Suomalainen, A. Gotz, T. Suortti, H. Yki-Jarvinen, M. Oresic, J. Kaprio, L. Peltonen, *Plos Med* 2008, 5.
- [54] L. Xiao, X. Yang, Y. Lin, S. Li, J. Jiang, S. Qian, Q. Tang, R. He, X. Li, *Int J Obesity* 2016, 40, 112.
- [55] J. Phieler, R. Garcia-Martin, J. D. Lambris, T. Chavakis, *Semin Immunol* 2013, 25, 47.
- [56] C. M. Phillips, E. Kesse-Guyot, N. Ahluwalia, R. McManus, S. Hercberg, D. Lairon, R. Planells, H. M. Roche, *Atherosclerosis* 2012, 220, 513.
- [57] J. N. Wei, H. Y. Li, F. C. Sung, C. C. Lin, C. C. Chiang, A. M. Carter, L. M. Chuang, *Pediatr Diabetes* 2012, 13, 476.
- [58] N. M. Al-Daghri, O. S. Al-Attas, H. E. Johnston, A. Singhanian, M. S. Alokail, K. M. Alkharfy, S. H. Abd-Alrahman, S. I. Sabico, T. I. Roumeliotis, A. Manousopoulou-Garbis, P. A. Townsend, C. H. Woelk, G. P. Chrousos, S. D. Garbis, *J Proteome Res* 2014, 13, 5094.
- [59] L. Milis, C. A. Morris, M. C. Sheehan, J. A. Charlesworth, B. A. Pussell, *Clin Exp Immunol* 1993, 92, 114.
- [60] O. Cominetti, A. N. Galindo, J. Cortesy, S. O. Moreno, I. Irincheeva, A. Valsesia, A. Astrup, W. H. M. Saris, J. Hager, M. Kussmann, L. Dayon, *J Proteome Res* 2016, 15, 389.
- [61] J. Esparza-Gordillo, J. M. Soria, A. Buil, L. Almasy, J. Blangero, J. Fontcuberta, S. R. de Cordoba, *Immunogenetics* 2004, 56, 77.